VERSION* | VERSION**

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/705, 16/28, C12N 15/11, A61K 39/00, 31/70, C12N 1/21, 5/10

A1

(11) International Publication Number: WO 97/42321

(43) International Publication Date: 13 November 1997 (13.11.97)

(21) International Application Number: PCT/US97/07856

(22) International Filing Date: 9 May 1997 (09.05.97)

(30) Priority Data:
08/647,397 9 May 1996 (09.05.96) US

(71) Applicant *(for all designated States except US)*: BRIGHAM & WOMEN'S HOSPITAL, INC. [US/US]; 75 Francis Street, Boston, MA 02115 (US).

(72) Inventors; and

(75) Inventors/Applican *(for US only)*: BEIER, David, R. [US/US]; 39 Arlington Road, Brookline, MA 02146 (US). BRADY, Kevin, P. [US/US]; 3 N: Bennet Court, Boston, MA 02113 (US).

(74) Agent: GATES, Edward, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

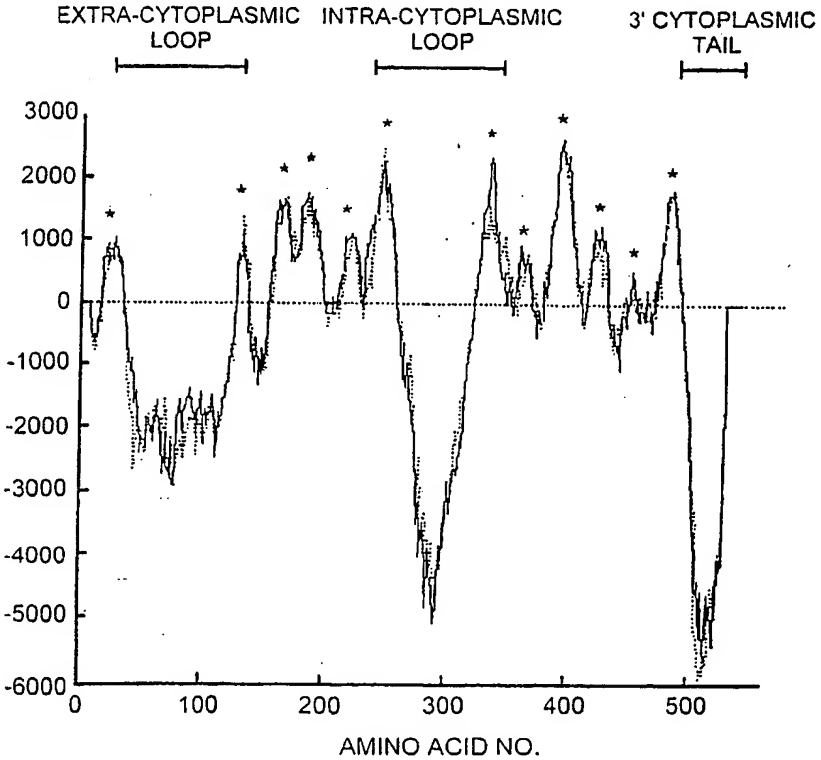
(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: OSTEOCLAST TRANSPORTER PROTEIN

(57) Abstract

The invention pertains to osteoclast transporter proteins and genes encoding such proteins, as well as fragments and biologically functional variants thereof. The invention also pertains to therapeutics and diagnostics involving the foregoing proteins and genes and agents that bind the foregoing proteins and genes, to influence osteoclast activity and bone remodeling.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

OSTEOCLAST TRANSPORTER PROTEIN

Government Support

This work was funded in part by the NIDDK under Grant Number R01 DK45639. The United States government may have certain rights to this invention.

Field of the Invention

This invention relates to osteoclast transporter genes, proteins coded for by such genes, and diagnostics and therapeutics related to medical conditions associated with such genes and proteins, including osteoporosis and osteopetrosis.

Background of the Invention

The remodeling of bone is a dynamic process. Cells continuously lay down and resorb bone material. An imbalance in the activity of cells that lay down new bone (osteoblasts) and cells that resorb bone (osteoclasts) can result in serious, and sometimes even fatal, disorders.

Osteoporosis is a term used for a number of diseases of diverse etiology, all involving a *reduction* in the mass of bone per unit volume. Osteoporosis is the most common of the metabolic bone diseases. Twenty-five million people in the United States and more than two hundred million people worldwide are affected by osteoporosis. Osteoporosis is frequent among post-menopausal women and is an important cause of morbidity in the elderly. It commonly results in bone fractures, and death can be a frequent occurance in the months following fractures, particularly those of the hip in elderly individuals.

Osteopetrosis is a disorder involving an *increase* in the mass of bone per unit volume. Its incidence is rare compared to osteoporosis, but it typically is life threatening. Despite having multiple causes, a defect in bone resorption is always the underlying mechanism. In many instances, the disorder is inherited as an autosomal recessive trait and involves abnormal osteoclast function. Bone marrow transplants from normal donors have been attempted to restore normal osteoclast precursor cells, but this therapy has shown only limited success.

Present treatments for osteoporosis and osteopetrosis are inadequate.

There exists a need to influence favorably the bone remodeling process to treat osteoporosis and osteopetrosis. There also exists a need to identify the gene(s) responsible for osteopetrosis and to provide a genetic therapy for treating osteopetrosis.

An object of the invention is to provide compounds that desirably influence the bone remodeling process.

Another object of the invention is to provide therapeutics for treating osteoporosis.

Another object of the invention is to provide therapeutics for treating osteopetrosis.

Still another object of the invention is to provide diagnostics and research tools relating to osteoporosis and osteopetrosis. These and other objects will be described in greater detail below.

Summary of the Invention

The invention involves in one respect the discovery of osteoclast transporter proteins and nucleic acid molecules encoding those proteins. The expression and biological activity of the proteins are necessary for normal osteoclast function, and alteration of the expression or biological activity of these proteins can be used to influence osteoclast activity and thereby affect bone remodeling. In addition, normal osteoclast function can be established in abnormal osteoclasts lacking a normal osteoclast transporter protein by supplying to the abnormal osteoclast a nucleic acid expressing a functional osteoclast transporter protein.

The preferred nucleic acids of the invention are homologs and alleles of the nucleic acids of SEQ ID NO:1 and SEQ ID NO:3. The invention further embraces functional equivalents, variants, analogs and fragments of the foregoing nucleic acids and also embraces proteins and peptides coded for by any of the foregoing.

According to one aspect of the invention, an isolated nucleic acid molecule is provided. The molecule hybridizes under stringent conditions to a molecule consisting of the nucleic acid sequence of SEQ ID NO:1 or SEQ ID NO:7 and it codes for an osteoclast transporter molecule. The invention further embraces nucleic acid molecules that differ from the foregoing isolated nucleic acid molecules in codon sequence due to the degeneracy of the genetic code. The invention also embraces complements of the foregoing nucleic acids.

Preferred isolated nucleic acid molecules are those which hybridize under stringent conditions to a nucleic acid molecule consisting of the nucleotide sequence of SEQ ID NO:3, particularly those comprising the human cDNAs or gene corresponding to SEQ ID NO:3 and the cDNAs or gene corresponding to SEQ ID NO:1.

The invention in another aspect involves expression vectors, and host cells transformed or transfected with such expression vectors, comprising the nucleic acid molecules described above.

In one embodiment of the invention, the host cell is a hematopoietic osteoclast precursor cell, such as a stem cell.

According to another aspect of the invention, an isolated nucleic acid molecule is provided which comprises a unique fragment of SEQ ID NO:1 between 12 and 2101 nucleotides in length or a unique fragment of nucleotides 1-2112 of SEQ ID NO:3 between 12 and 2111 nucleotides in length, and complements thereof. In one embodiment, the unique fragment is at least 150 and more preferably at least 200 nucleotides in length. In another embodiment, the unique fragment is between 12 and 32 contiguous nucleotides in length. In some embodiments, the isolated nucleic acid molecule is a unique fragment of SEQ ID NO:1 between 12 and 1974 nucleotides in length or a unique fragment of nucleotides 120-1733 of SEQ ID NO:1 between 12 and 1612 nucleotides in length. In another embodiment, the isolated nucleic acid molecule is selected from the group consisting of a unique fragment of nucleotides 1-1179 of SEQ ID NO:3 between 12 and 1178 nucleotides in length and a unique fragment of nucleotides 1512-2112 of SEQ ID NO:3 between 12 and 599 nucleotides in length. In certain embodiments, the isolated nucleic acid molecule is a unique fragment of the coding region of the human osteoclast transporter gene. Such isolated nucleic acid molecule include unique fragments of nucleotides 73-1758 of SEQ ID NO:3 between 12 and 1684 nucleotides in length. Preferably, unique fragments of the coding region include isolated nucleic acid molecules selected from the group consisting of unique fragments of nucleotides 73-1179 of SEQ ID NO:3 between 12 and 1105 nucleotides in length and unique fragments of nucleotides 1512-1758 of SEQ ID NO:3 between 12 and 245 nucleotides in length.

According to another aspect of the invention, isolated polypeptides coded for by the isolated nucleic acid molecules described above also are provided as well as functional equivalents, variants, analogs and fragments thereof. In one embodiment, the polypeptide is a human osteoclast transporter protein or a functionally active fragment or variants thereof.

The invention also provides isolated polypeptides which selectively bind an osteoclast transporter protein or fragments thereof. Isolated binding polypeptides include antibodies and fragments of antibodies (e.g. Fab, $F(ab)_2$, Fd and antibody fragments which include a CDR3 region which binds selectively to the osteoclast transporter proteins of the invention). Preferred isolated binding polypeptides are those that bind an extracellular portion of the osteoclast transporter proteins of the invention.

The invention in another aspect involves a method for decreasing osteoclast activity in a subject. An agent that selectively binds to an isolated nucleic acid molecule as described above or an expression product thereof is administered to a subject in need of such treatment, in an amount effective to decrease osteoclast activity in the subject. In one embodiment, the agent selectively binds to an extracellular domain of the expression product. Preferred agents are modified antisense nucleic acids and polypeptides.

The invention is a further aspect provides for the use of agents which bind to the foregoing nucleic acids and expression products thereof in the preparation of a medicament. Preferred agents are modified antisense nucleic acids and polypeptides.

The invention also contemplates gene therapy for osteopetrosis, wherein defective stem cells of a donor are genetically engineered to include an isolated nucleic acid expressing a functional osteoclast transporter protein. The cells then are returned to the donor.

In the same manner as described above kidney cell tubule activity may be modulated using the methods and products of the invention.

Brief Description of the Drawing

FIG. 1 is a graph depicting hydrophobicity vs. amino acid for the mouse osteoclast transporter protein.

Detailed Description of the Invention

The present invention in one aspect involves the cloning of a gene encoding an osteoclast transporter protein. The sequence of the gene (from mouse) is presented as SEQ ID NO:1. and the predicted amino acid sequence of this gene's protein product is presented as SEQ ID NO:2. The mRNA transcript is about 2.0 kb in length. It was obtained from mouse kidney tubule cells. A partial clone of the apparent human homolog was identified in GenBank, accession no. H20345, and is presented as SEQ ID NO:7. The partial clone is available from American Type Culture Collection Depository. Rockville, Maryland. The human cDNA encoding the osteoclast transporter protein was isolated by a standard hybridization protocol using probes consisting of the 5' end of the mouse gene (nucleotides 45-357 of SEQ ID NO:1) and the partial human gene sequence (SEQ ID NO:7). The sequence of the full length human cDNA is presented as SEQ ID NO:3, and the predicted amino acid sequence of the human protein product is presented as SEQ ID NO:4.

The mouse gene maps to the proximal portion of chromosome 19, in a region to which the osteosclerosis (*oc*) mutation has been assigned previously. The phenotype of *oc/oc* mutant mice includes osteopetrosis, and the osteoclasts of these mutant mice, although present, appear to be nonfunctional. The osteoclast transporter protein of the present invention is not expressed in oc/oc mutant mice, although Southern analysis indicates that the transporter locus is unrearranged and intact in these mutant mice. The osteoclast transporter protein of the present invention, however, is expressed in normal osteoclasts, and it is believed that hereditary osteopetrosis results from mutations to this gene, which impair osteoclast function.

The gene is not expressed in any other normal tissue tested, except in kidney tubule cells where its function is believed to be redundant. In particular, Northern analysis demonstrated that the gene is not expressed in normal skeletal muscle, pancreas, lung, heart, brain, stomach, spleen, salivary gland, thymus, liver, large intestine or small intestine tissue.

Searches of GenBank for similar related proteins shows that this protein shares some very limited, localized homology and sequence motifs to known proteins with transport functions. The protein of the invention, however, does have a length and tertiary structures similar to known transporters.

Specifically, it appears to be a member of a large family of genes with transport functions called the "major facilitator superfamily" by Marger and Saier, which include a number of uni-, sym- and antiporters specific for sugars, organic acids and drugs. The common structural motif of this family are 12 transmembrane alpha helices, and often includes a central cytoplasmic loop between the 6th and 7th membrane-spanning domains. The hydrophobicity plot for the protein of the present invention is shown in Fig. 1. Each asterisk indicates a potential membrane spanning hydrophobic domain. There appears to be an extracellular loop between the first and second hydrophobic domains, an intra-cytoplasmic loop between the hydrophobic domains 6 and 7, and a 3' intra-cytoplasmic tail. The 5' end also appears to be internal within the cytoplasm.

The transporter also contains amino-acid motifs found in this class of molecules, including a sequence, D-R-F-G-R-K (SEQ ID NO:5), similar to a D-R/K-X-R-R/K sequence (SEQ ID NO:6) after the 2nd membrane domain, and a P-E-S/T sequence found after the 12th transmembrane domain. For genes with known function, the transporter is most closely related to the polyspecific rat cation transporter Octl (30% amino acid identity and 50% similarity), which has been shown to transport a variety of organic cationic drugs (Grundemann et al., *Nature*, Vol. 372, December 1994.)

The invention thus involves osteoclast transport proteins, genes encoding those proteins, functional modifications and variants of the foregoing, useful fragments of the foregoing, as well as therapeutics and diagnostics relating thereto.

Homologs and alleles of the osteoclast transporter genes of the invention can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for osteoclast transporter proteins and which hybridize to a nucleic acid molecule consisting of SEQ ID NO:1 or SEQ ID NO:3, under stringent conditions. The term "stringent conditions" as used herein refers to parameters with which the art is familiar. More specifically, stringent conditions, as used herein, refers to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH_2PO_4(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, ph7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed at 2 x SSC at room temperature and then at 0.1 x SSC/0.1 x SDS at 65°C.

There are other conditions, reagents, and so forth which can used, which result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of osteoclast transporter proteins of the invention. The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

In general homologs and alleles typically will share at least 40% nucleotide identity and/or at least 50% amino acid identity to SEQ ID NOs:1 or 3 and SEQ ID NOs:2 or 4, respectively, in some instances will share at least 50% nucleotide identity and/or at least 65% amino acid identity and in still other instances will share at least 60% nucleotide identity and/or at least 75% amino acid identity. Watson-Crick complements of the forgoing nucleic acids also are embraced by the invention.

In screening for osteoclast transporter protein family members, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against x-ray film to detect the radioactive signal.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, *in vitro* or *in vivo*, to incorporate a serine residue. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

The invention also provides isolated unique fragments of SEQ ID NO:1, SEQ ID NO:3, or compliments of SEQ ID NO:1 or SEQ ID NO:3. A unique fragment is one that is a 'signature' for the larger nucleic acid. It, for example, is long enough to assure that its precise sequence is not found in molecules outside of the osteoclast transporter protein family as defined by claim 1. Unique fragments can be used as probes in Southern blot assays to identify family members or can be used in amplification assays such as those employing PCR. As known to those skilled in the art, large probes such as 200 nucleotides or more are preferred for certain uses such as Southern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins for generating antibodies or for generating immunoassay components. Likewise, unique fragments can be employed to produce fragments of the osteoclast transporter protein such as only the extracellular portion, useful, for example, in immunoassays or as a competitive inhibitor of the substrate of the osteoclast transporter protein in therapeutic applications. Unique fragments further can be used as antisense molecules to inhibit the expression of the osteoclast transporter proteins of the invention, particularly for therapeutic purposes as described in greater detail below.

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of SEQ ID NO:1 and SEQ ID NO:3, and their complements, will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides (e.g. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 bases long). Virtually any segment of

SEQ ID NO:1 or SEQ ID NO:3, or their complements, that is 18 or more nucleotides in length will be unique. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-family members. A comparison of the sequence of the fragment to those on known data bases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding an osteoclast transporter protein, to decrease osteoclast activity. This is desirable in virtually any medical condition wherein a reduction in osteoclast activity is desirable, including when a reduction in bone loss or an increase in bone mass is desired. By decreasing the osteoclast activity in a subject, bone remodeling thus can be favorably affected. Antisense molecules, in this manner, can be used to slow down or arrest the loss in bone mass occurring with certain forms of osteoporosis and may even result in an increase in bone mass in circumstances where an increase in bone mass is desirable.

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon SEQ ID NO:1 and SEQ ID NO:3, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target. Most preferably, the antisense

oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In addition, 3'-untranslated regions or telomerase sites may be targeted. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., Cell Mol. Neurobiol. 14(5):439-457 (1994)) and at which proteins are not expected to bind. Finally, although, SEQ ID NOs:1 and 3 disclose cDNA sequences, one of ordinary skill in the art may easily derive the genomic DNAs corresponding to the cDNAs of SEQ ID NOs:1 and 3. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNAs corresponding to SEQ ID NOs:1 and 3. Similarly, antisense to allelic or homologous cDNAs and genomic DNAs are enabled without undue experimentation.

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer.

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred synthetic internucleoside linkages include phosphorothioates, alkylphosphonates, phosphorodithioates, phosphate esters, alkylphosphonothioates, phosphoramidates, carbamates, carbonates, phosphate triesters, acetamidates, and carboxymethyl esters.

The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-O-alkylated ribose group. In addition, modified oligonucleotides may include sugars such as arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding osteoclast transporter proteins, together with pharmaceutically acceptable carriers.

Antisense oligonucleotides may be administered as part of a pharmaceutical composition or medicament. Such a pharmaceutical composition or medicament may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

The invention also involves expression vectors coding for osteoclast transporter proteins and fragments and variants thereof and host cells containing those expression vectors. Virtually any cells, prokaryotic or eukaryotic, which can be transformed with heterologous DNA or RNA and which can be grown or maintained in culture, may be used in the practice of the invention. Examples include bacterial cells such as *E. coli* and mammalian cells such as mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, including mast cells, fibroblasts, oocytes and lymphocytes, and they may be primary cells or cell lines. Specific examples include CHO cells and COS cells. Cell-free transcription systems also may be used in lieu of cells. In gene therapy applications, human hematopoietic cells that are precursors of osteoclasts are contemplated.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to, plasmids and phagemids. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g. ß-galactosidase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hots, colonies or plaques. Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were

capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribing and 5' non-translating sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribing regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences, 5' or 3'. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See Sanbrook et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding the osteoclast transporter protein or fragment or variant thereof. That heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell. In still another aspect of the invention, a defective osteoclast or precursor thereof is treated with DNA in a manner to promote via homologous recombination intracellularly the correction of a defective osteoclast transporter gene.

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV (available from Invitrogen, San Diego, CA) that contain a selectable marker such as a gene that confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen), which contains an Epstein Barr virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element.

mRNA expression and transporter function can be tested using these vectors in a wide variety of mammalian cell lines. Preferred systems include cells derived from kidney tubules, including MDCK and IMCD-3 cells (both available from ATCC).

A variety of systems for expression of proteins in bacterial, yeast, mammalian, or insect cells have been described and are commercially available. Preferred systems include the

Glutathione-S-transferase (GST) Gene Fusion system available from Pharmacia Biotech, Piscataway, NJ. In this system a plasmid is constructed containing the protein sequence of interest (in this case, the transporter including the first extracellular domain) inserted in frame downstream of the 25 kDa GST domain from S. japonicum. Expression of the fusion protein can be induced in transfected bacterial cells and the fusion protein purified by affinity chromatography using Glutathione Sepharose 4B. Cleavage of the desired peptide from the GST sequences is achieved using a site specific protease whose recognition sequence is located immediately upstream from the cloning site. An alternative system which is desirable since it maintains eucaryotic-specific functions such as glycosylation is recombination into baculovirus. Standard protocols exist (c.f. O'Reilly et al., Baculovirus Expression Vectors: A :Laboratory Manual, IRL/Oxford University Press, 1992) and vectors, cells, and reagents are commercially available.

The invention also involves polypeptides which bind to osteoclast transporter proteins and in certain embodiments preferably to the extracellular domain of the proteins. Such binding partners can be used in screening assays to detect the presence or absence of the osteoclast transporter protein and in purification protocols to isolate osteoclast transporter proteins. Such binding partners also can be used to inhibit the native activity of the osteoclast transporter protein by binding to the extracellular domain of such proteins. Likewise, such binding partners can be used to selectively target drugs, toxins or other molecules to osteoclasts. In this manner, osteoclast activity may be selectively increased by drugs that would enhance osteoclast activity or selectively impaired by drugs that would inhibit osteoclast activity (e.g. toxins, antisense, etc.). In this manner, bone remodeling may be desirably affected.

The invention, therefore, involves antibodies or fragments of antibodies having the ability to selectively bind to osteoclast transporter proteins, and preferably to the extracellular domain thereof. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology. Antibodies were raised according to standard procedures. A fusion protein was prepared using the pGEX-4T-2 vector (Pharmacia, Piscataway, NJ). The fusion protein contains a glutathione-S-transferase (GST) fragment joined to the terminal 84 amino acids of the osteoclast transporter. The fusion protein was expressed and injected into rabbits to produce antibodies as rabbit immune serum. The antibodies so generated detected a protein of molecular weight which corresponds to the osteoclast transporter in tissue prepared from normal mouse kidneys. The protein is significantly reduced or absent in kidney tissue prepared from

oc/oc mutant mice, consistent with the absence of detectable osteoclast transporter mRNA by Northern blot in these mice.

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated an $F(ab')_2$ fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have

been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for $F(ab')_2$, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric $F(ab')_2$ fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies. Thus, the invention involves polypeptides of numerous size and type that bind specifically to osteoclast transporter proteins. These polypeptides may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptoids and non-peptide synthetic moieties.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the osteoclast transporter protein. This process can be repeated through several cycles of reselection of phage that bind to the osteoclast transporter protein. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear portion of the sequence that binds to the osteoclast transporter protein can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Thus, the osteoclast transporter molecule of the invention, an extracellular domain thereof, or the like, can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding partners of the extracellular portion of the osteoclast

transporter protein of the invention. Such molecules can be used, as described, for screening assays, for diagnostic assays, for purification protocols, for interfering directly with the functioning of osteoclasts by binding to the osteoclast transporter protein or for targeting drugs, toxins and/or labeling agents (e.g. radioisotopes, fluorescent molecules, etc.) to osteoclasts and/or osteoclast transporter proteins. Drug molecules that would affect osteoclast activity and toxin molecules that would disable or destroy osteoclast cells are known to those skilled in the art and are commercially available. For example, the immunotoxin art provides examples of toxins which are effective when delivered to a cell by an antibody or fragment thereof. Examples of toxins include ribosome-damaging toxins derived from plants or bacterial such as ricin, abrin, saporin, Pseudomonas endotoxin, diphtheria toxin, A chain toxins, blocked ricin, etc.

Osteoclast activity can be assayed both in vitro and in vivo. Heterogeneous culture systems in which bone marrow derived cells are co-cultured with osteoblast cells in the presence of 1,25-dihydroxyvitamin D3 will generate mature osteoclasts, whose activity can be quantitated by measuring their ability to form resorption pits on dentine slices (Takahashi et al., *Endocrinology*, Vol. 122, No. 4, 1988). In mammals, decreased osteoclast function results in the clinical condition of osteopetrosis. In rodents, this can be readily scored by virtue of the fact that, in this case, incisors fair to erupt after birth. Thus, rodent neonates can be treated with compositions of the invention and the effect on osteoclast activity can be determined by scoring incisor eruption. The degree of osteopetrosis can be more accurately quantitated by morphometrical analysis of the bone marrow space (Boyce et al., *J. Clin. Invest.*, Vol. 90, 1992).

When used therapeutically, the compounds of the invention are administered in therapeutically effective amounts. In general, a therapeutically effective amount means that amount necessary to delay the onset of, inhibit the progression of, or halt altogether the particular condition being treated. Therapeutically effective amounts specifically will be those which desirably influence osteoclast activity. When it is desired to decrease osteoclast activity, then any inhibition of osteoclast activity is regarded as a therapeutically effective amount. When it is desired to increase osteoclast activity, then any enhancement of osteoclast activity is regarded as a therapeutically effective amount. Generally, a therapeutically effective amount will vary with the subject's age, condition, and sex, as well as the nature and extent of the disease in the subject, all of which can be determined by one of ordinary skill in the art. The dosage may be adjusted by the individual physician or veterinarian, particularly in the event of any complication. A therapeutically effective amount typically varies from 0.01 mg/kg to about 1000 mg/kg,

preferably from about 0.1 mg/kg to about 200 mg/kg and most preferably from about 0.2 mg//kg to about 20 mg/kg, in one or more dose administrations daily, for one or more days.

The therapeutics of the invention can be administered by any conventional route, including injection or by gradual infusion over time. The administration may, for example, be oral, intravenous, intraperitoneal, intramuscular, intracavity, subcutaneous, or transdermal. When antibodies are used therapeutically, a preferred route of administration is by pulmonary aerosol. Techniques for preparing aerosol delivery systems containing antibodies are well known to those of skill in the art. Generally, such systems should utilize components which will not significantly impair the biological properties of the antibodies, such as the paratope binding capacity (see, for example, Sciarra and Cutie, "Aerosols," in Remington's Pharmaceutical Sciences, 18th edition, 1990, pp 1694-1712; incorporated by reference). Those of skill in the art can readily determine the various parameters and conditions for producing antibody aerosols without resort to undue experimentation. When using antisense preparations of the invention, slow intravenous administration is preferred.

Preparations for parenteral administration include sterile aqueous or non-aqueous solutions. suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives may also be present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases and the like.

The invention also contemplates gene therapy. The procedure for performing ex vivo gene therapy is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction in vitro of a functional copy of a gene into a cell(s) of a subject which contains a defective copy of the gene, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which

are described in PCT application WO95/00654. In vivo gene therapy using vectors such as adenovirus also is contemplated according to the invention.

Thus, defective osteoclasts or precursors thereof are provided with a non-defective nucleic acid encoding an osteoclast transporter protein. In particular, such gene therapy is appropriate for hereditary forms of osteopetrosis involving defective osteoclast transporter protein genes. For example, primary human blood cells which are precursors of osteoclasts can be obtained from the bone marrow of a subject who is a candidate for such gene therapy. Candidates can be identified by screening for abnormal osteoclast function that results from a defective osteoclast transporter protein. Then, such cells can be genetically engineered *ex vivo* with DNA (RNA) encoding a normal osteoclast transporter protein. The genetically engineered cells then are returned to the patient.

That the transporter can be used to correct the defect in the clinical conditions of osteopetrosis in which it has been inactivated is demonstrated using a mouse model system. The osteosclerosis (*oc*) mutant mouse fails to express the transporter and has osteopetrosis. The transporter is introduced into a vector such as pXT1 (Strategene, La Jolla, CA) that, when transfected into an appropriate packaging cell line, will generate a higher titer of replication-defective retrovitral particles carrying the transporter under the regulatory control of a thymidine kinase promoter. Hematopoietic progenitor cells from *oc* mutant mice are prepared and, using well-documented techniques of infection and selection, stem cells carrying this vector are selected after coculture with the retrovirus. Re-population of lethally irradiated *oc* mutant mice with this transfected population is accomplished using documented protocols of hematopoietic transplantation, and the degree of amelioration of the osteopetrosis phenotype then is determined by radiography and morphometrical analysis.

While the invention has been described with respect to certain embodiments, it should be appreciated that many modifications and changes may be made by those of ordinary skill in the art without departing from the spirit of the invention. It is intended that such modification, changes and equivalents fall within the scope of the following claims.

A Sequence Listing is provided, followed by what is claimed:

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: BRIGHAM AND WOMEN'S HOSPITAL, INC.

(B) STREET: 75 FRANCIS STREET

(C) CITY: BOSTON

(D) STATE: MASSACHUSETTS

(E) COUNTRY: UNITED STATES OF AMERICA

(F) POSTAL CODE: 02115

(i) APPLICANT/INVENTOR:

(A) NAME: BEIER, DAVID R.

(B) STREET: 39 ARLINGTON ROAD

(C) CITY: BROOKLINE

(D) STATE: MASSACHUSETTS

(E) COUNTRY: UNITED STATES OF AMERICA

(F) POSTAL CODE: 02146

(i) APPLICANT/INVENTOR:

(A) NAME: BRADY, KEVIN P.

(B) STREET: 3N BENNET COURT

(C) CITY: BOSTON

(D) STATE: MASSACHUSETTS

(E) COUNTRY: UNITED STATES OF AMERICA

(F) POSTAL CODE: 02113

(ii) TITLE OF INVENTION: OSTEOCLAST TRANSPORTER

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.

(B) STREET: 600 ATLANTIC AVENUE

(C) CITY: BOSTON

(D) STATE: MASSACHUSETTS

(E) COUNTRY: UNITED STATES OF AMERICA

(F) POSTAL CODE: 02210

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/647,397

(B) FILING DATE: 09-MAY-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Gates, Edward R.

(B) REGISTRATION NUMBER: 31,616

(C) REFERENCE/DOCKET NUMBER: B0801/7048WO

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 617-720-3500

(B) TELEFAX: 617-720-2441

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2102 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mus musculus

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 120..1733

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TGCGAGCGTG CTACTACAGC AGCTGCTGAA CCTAGACAGG CACGGCAACT GCTGCATCCA 60

GCTCCAGCCC AACTGAATCC AGCTCCAACC ACCAGTTTTG GTTCATCTTG CCTGGTGCC 119

ATG ACC TTC TCC GAG ATT CTG GAC CGT GTT GGA AGC ATG GGC CCC TTC 167
Met Thr Phe Ser Glu Ile Leu Asp Arg Val Gly Ser Met Gly Pro Phe
1 5 10 15

CAG TAC CTG CAT GTG ACC TTG CTG GCC CTC CCA ATC CTC GGA ATA GCC 215

```
Gln Tyr Leu His Val Thr Leu Leu Ala Leu Pro Ile Leu Gly Ile Ala
                20                  25                  30

AAC CAC AAC TTG CTA CAG ATC TTC ACA GCC ACC ACC CCT GAC CAC CAC      263
Asn His Asn Leu Leu Gln Ile Phe Thr Ala Thr Thr Pro Asp His His
        35                  40                  45

TGT CGC CCG CCC CCC AAC GCC TCT CTA GAG CCC TGG GTA CTC CCC TTG      311
Cys Arg Pro Pro Pro Asn Ala Ser Leu Glu Pro Trp Val Leu Pro Leu
    50                  55                  60

GGC CCA AAC GGG AAG CCT GAG AAG TGT CTC CGC TTC GTG CAT CTG CCA      359
Gly Pro Asn Gly Lys Pro Glu Lys Cys Leu Arg Phe Val His Leu Pro
65                  70                  75                  80

AAC GCC AGT CTT CCC AAT GAC ACC CAG GGG GCC ACC GAG CCA TGC TTG      407
Asn Ala Ser Leu Pro Asn Asp Thr Gln Gly Ala Thr Glu Pro Cys Leu
                85                  90                  95

GAT GGC TGG ATC TAC AAC AGC ACC AGA GAC ACC ATT GTG ACA GAG TGG      455
Asp Gly Trp Ile Tyr Asn Ser Thr Arg Asp Thr Ile Val Thr Glu Trp
            100                 105                 110

GAC TTG GTA TGC GGC TCC AAC AAA CTG AAG GAG ATG GCA CAG TCA GTC      503
Asp Leu Val Cys Gly Ser Asn Lys Leu Lys Glu Met Ala Gln Ser Val
        115                 120                 125

TTC ATG GCA GGT ATA CTG GTT GGA GGA CCT GTG TTT GGA GAA CTG TCA      551
Phe Met Ala Gly Ile Leu Val Gly Gly Pro Val Phe Gly Glu Leu Ser
    130                 135                 140

GAC AGG TTT GGC CGC AAG CCC ATC CTG ACC TGG AGC TAT CTC TTG CTG      599
```

Asp Arg Phe Gly Arg Lys Pro Ile Leu Thr Trp Ser Tyr Leu Leu Leu
145 150 155 160

GCA GCC AGT GGC TCC AGT GCT GCC TTC AGC CCC AGC CTC ACT GTC TAT 647
Ala Ala Ser Gly Ser Ser Ala Ala Phe Ser Pro Ser Leu Thr Val Tyr
165 170 175

ATG ATC TTC CGA TTC CTG TGT GGC TGC AGC ATC TCG GGC ATT TCT CTG 695
Met Ile Phe Arg Phe Leu Cys Gly Cys Ser Ile Ser Gly Ile Ser Leu
180 185 190

AGC ACC ATT ATC TTG AAT GTG GAA TGG GTA CCC ACC TCC ACG CGG GCC 743
Ser Thr Ile Ile Leu Asn Val Glu Trp Val Pro Thr Ser Thr Arg Ala
195 200 205

ATC TCA TCA ACA ACT ATT GGG TAC TGC TAC ACC ATT GGT CAA TTC ATT 791
Ile Ser Ser Thr Thr Ile Gly Tyr Cys Tyr Thr Ile Gly Gln Phe Ile
210 215 220

CTG CCT GGC CTG GCC TAT GCC GTT CCT CAG TGG CGC TGG CTA CAG TTG 839
Leu Pro Gly Leu Ala Tyr Ala Val Pro Gln Trp Arg Trp Leu Gln Leu
225 230 235 240

TCC GTG TCT GCT GCC TTC TTC ATC TTC TCC TTG TTG TCC TGG TGG GTA 887
Ser Val Ser Ala Ala Phe Phe Ile Phe Ser Leu Leu Ser Trp Trp Val
245 250 255

CCA GAG TCC ATA CGC TGG CTG GTT CTG TCT GGA AAA TTC TCA CGA GCT 935
Pro Glu Ser Ile Arg Trp Leu Val Leu Ser Gly Lys Phe Ser Arg Ala
260 265 270

CTG AAG ACA CTC CAA CGT GTG GCT ACC TTC AAC GGC AAG AAG GAG GAA 983

Leu Lys Thr Leu Gln Arg Val Ala Thr Phe Asn Gly Lys Lys Glu Glu
275 280 285

GGG GAA AAG CTC ACT GTG GAG GAG CTG AAG TTC AAC TTG CAG AAG GAC 1031
Gly Glu Lys Leu Thr Val Glu Glu Leu Lys Phe Asn Leu Gln Lys Asp
290 295 300

ATC ACC TCA GCC AAG GTC AAA TAT GGC TTA TCT GAC TTG TTC CGA GTG 1079
Ile Thr Ser Ala Lys Val Lys Tyr Gly Leu Ser Asp Leu Phe Arg Val
305 310 315 320

TCC ATC CTG CGC CGT GTG ACC TTC TGT CTC TCT CTG GCC TGG TTT GCT 1127
Ser Ile Leu Arg Arg Val Thr Phe Cys Leu Ser Leu Ala Trp Phe Ala
325 330 335

ACT GGC TTT GCC TAC TAC AGT TTG GCT ATG GGA GTA GAA GAA TTT GGA 1175
Thr Gly Phe Ala Tyr Tyr Ser Leu Ala Met Gly Val Glu Glu Phe Gly
340 345 350

GTC AAC ATC TAC ATA CTC CAG ATC ATC TTC GGT GGG GTT GAC ATT CCC 1223
Val Asn Ile Tyr Ile Leu Gln Ile Ile Phe Gly Gly Val Asp Ile Pro
355 360 365

GCC AAG TTC ATC ACA ATC CTC TCC ATA AGT TAT CTG GGC CGG CGC ATC 1271
Ala Lys Phe Ile Thr Ile Leu Ser Ile Ser Tyr Leu Gly Arg Arg Ile
370 375 380

ACT CAG GGC TTC CTC CTG ATC CTG GCA GGA GTG GCC ATC CTG GCC CTC 1319
Thr Gln Gly Phe Leu Leu Ile Leu Ala Gly Val Ala Ile Leu Ala Leu
385 390 395 400

ATC TTT GTG TCT TCA GAA ATG CAG CTC TTG AGA ACA GCA CTG GCT GTA 1367

Ile Phe Val Ser Ser Glu Met Gln Leu Leu Arg Thr Ala Leu Ala Val
405 410 415

TTT GGG AAG GGA TGC CTG TCT GGC TCC TTC AGC TGC CTC TTC CTC TAC 1415
Phe Gly Lys Gly Cys Leu Ser Gly Ser Phe Ser Cys Leu Phe Leu Tyr
420 425 430

ACA AGT GAG CTC TAC CCT ACA GTC CTC AGG CAA ACA GGT ATG GGT ATC 1463
Thr Ser Glu Leu Tyr Pro Thr Val Leu Arg Gln Thr Gly Met Gly Ile
435 440 445

AGT AAC ATA TGG GCT CGA GTG GGA AGT ATG ATA GCC CCA CTG GTG AAA 1511
Ser Asn Ile Trp Ala Arg Val Gly Ser Met Ile Ala Pro Leu Val Lys
450 455 460

ATC ACG GGA GAA CTG CAG CCC TTC ATC CCT AAT GTC ATC TTT TGG ACC 1559
Ile Thr Gly Glu Leu Gln Pro Phe Ile Pro Asn Val Ile Phe Trp Thr
465 470 475 480

ATG ACT CTA CTG GGA GGC AGT GCT GCC TTC TTT CTG CTT GAG ACC CTC 1607
Met Thr Leu Leu Gly Gly Ser Ala Ala Phe Phe Leu Leu Glu Thr Leu
485 490 495

AAT CGG CCC TTA CCA GAA ACT ATC GAG GAC ATA CAA GAC TGG TAC CAG 1655
Asn Arg Pro Leu Pro Glu Thr Ile Glu Asp Ile Gln Asp Trp Tyr Gln
500 505 510

CAA ACC AAG AAA ACA AAG CAG GAG CCA GAA GCA GAA AAG GCA TCC CAG 1703
Gln Thr Lys Lys Thr Lys Gln Glu Pro Glu Ala Glu Lys Ala Ser Gln
515 520 525

ACA ATC CCG CTG AAG ACT GGT GGA CCC TAGCTAAGAA CAACAGAATC 1750

Thr Ile Pro Leu Lys Thr Gly Gly Pro
530 535

CTCTTTCCTG CCTTCCAGAG ACTGATCCCA AGCAGGGCCC TTCCAAGGCT ATTCGAGCAC 1810

CTTAGGGGTT GGGTGGAGCC CCAGCTGGCT CCATGCTCTC AGAACAAAGA CTTCTGAGAG 1870

TTCAGCAAAA GTGTTTTACC TTCACCACCT CCACCGTAGC CCACAACCCA GACCTGGCCT 1930

GTTCACAGCC CTAGCCATAC TCACTCCTGC ACTCATCCTC CCTGCAACCC AGGCCCTGCC 1990

ATTTTTCTCT ACCCTCTTTG TATTGGCCAT TTCCTCCATT GTCCCACCTC CATTTCCCTT 2050

TGAGATTCCC TGGCAGTTCT AATGGTTTCC TCTTACCCTC CCCCTCGTGC CG 2102

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 537 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Thr Phe Ser Glu Ile Leu Asp Arg Val Gly Ser Met Gly Pro Phe
1 5 10 15

Gln Tyr Leu His Val Thr Leu Leu Ala Leu Pro Ile Leu Gly Ile Ala
20 25 30

Asn His Asn Leu Leu Gln Ile Phe Thr Ala Thr Thr Pro Asp His His
35 40 45

Cys Arg Pro Pro Pro Asn Ala Ser Leu Glu Pro Trp Val Leu Pro Leu
50 55 60

Gly Pro Asn Gly Lys Pro Glu Lys Cys Leu Arg Phe Val His Leu Pro
65 70 75 80

Asn Ala Ser Leu Pro Asn Asp Thr Gln Gly Ala Thr Glu Pro Cys Leu
85 90 95

Asp Gly Trp Ile Tyr Asn Ser Thr Arg Asp Thr Ile Val Thr Glu Trp
100 105 110

Asp Leu Val Cys Gly Ser Asn Lys Leu Lys Glu Met Ala Gln Ser Val
115 120 125

Phe Met Ala Gly Ile Leu Val Gly Gly Pro Val Phe Gly Glu Leu Ser
130 135 140

Asp Arg Phe Gly Arg Lys Pro Ile Leu Thr Trp Ser Tyr Leu Leu Leu
145 150 155 160

Ala Ala Ser Gly Ser Ser Ala Ala Phe Ser Pro Ser Leu Thr Val Tyr
165 170 175

Met Ile Phe Arg Phe Leu Cys Gly Cys Ser Ile Ser Gly Ile Ser Leu
180 185 190

Ser Thr Ile Ile Leu Asn Val Glu Trp Val Pro Thr Ser Thr Arg Ala
195 200 205

Ile Ser Ser Thr Thr Ile Gly Tyr Cys Tyr Thr Ile Gly Gln Phe Ile
210 215 220

Leu Pro Gly Leu Ala Tyr Ala Val Pro Gln Trp Arg Trp Leu Gln Leu
225 230 235 240

Ser Val Ser Ala Ala Phe Phe Ile Phe Ser Leu Leu Ser Trp Trp Val
245 250 255

Pro Glu Ser Ile Arg Trp Leu Val Leu Ser Gly Lys Phe Ser Arg Ala
260 265 270

Leu Lys Thr Leu Gln Arg Val Ala Thr Phe Asn Gly Lys Lys Glu Glu
275 280 285

Gly Glu Lys Leu Thr Val Glu Glu Leu Lys Phe Asn Leu Gln Lys Asp
290 295 300

Ile Thr Ser Ala Lys Val Lys Tyr Gly Leu Ser Asp Leu Phe Arg Val
305 310 315 320

Ser Ile Leu Arg Arg Val Thr Phe Cys Leu Ser Leu Ala Trp Phe Ala
325 330 335

Thr Gly Phe Ala Tyr Tyr Ser Leu Ala Met Gly Val Glu Glu Phe Gly
340 345 350

Val Asn Ile Tyr Ile Leu Gln Ile Ile Phe Gly Gly Val Asp Ile Pro
355 360 365

Ala Lys Phe Ile Thr Ile Leu Ser Ile Ser Tyr Leu Gly Arg Arg Ile
370 375 380

Thr Gln Gly Phe Leu Leu Ile Leu Ala Gly Val Ala Ile Leu Ala Leu
385 390 395 400

Ile Phe Val Ser Ser Glu Met Gln Leu Leu Arg Thr Ala Leu Ala Val
405 410 415

Phe Gly Lys Gly Cys Leu Ser Gly Ser Phe Ser Cys Leu Phe Leu Tyr
420 425 430

Thr Ser Glu Leu Tyr Pro Thr Val Leu Arg Gln Thr Gly Met Gly Ile
435 440 445

Ser Asn Ile Trp Ala Arg Val Gly Ser Met Ile Ala Pro Leu Val Lys
450 455 460

Ile Thr Gly Glu Leu Gln Pro Phe Ile Pro Asn Val Ile Phe Trp Thr
465 470 475 480

Met Thr Leu Leu Gly Gly Ser Ala Ala Phe Phe Leu Leu Glu Thr Leu
485 490 495

Asn Arg Pro Leu Pro Glu Thr Ile Glu Asp Ile Gln Asp Trp Tyr Gln
500 505 510

Gln Thr Lys Lys Thr Lys Gln Glu Pro Glu Ala Glu Lys Ala Ser Gln
515 520 525

Thr Ile Pro Leu Lys Thr Gly Gly Pro
530 535

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2135 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 73..1758

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAATTCCGGT GATCACCAGC CCCATCGGAT CCAGACCCGG CCACCAGCTC TGGCTCGTCT 60

TGCCCCAGTG CC ATG ACC TTC TCG GAG ATC CTG GAC CGT GTG GGA AGC 108
Met Thr Phe Ser Glu Ile Leu Asp Arg Val Gly Ser
1 5 10

ATG GGC CAT TTC CAG TTC CTG CAT GTA GCC ATA CTG GGC CTC CCG ATC 156
Met Gly His Phe Gln Phe Leu His Val Ala Ile Leu Gly Leu Pro Ile
15 20 25

CTC AAC ATG GCC AAC CAC AAC CTG CTG CAG ATC TTC ACA GCC GCC ACC 204

Leu Asn Met Ala Asn His Asn Leu Leu Gln Ile Phe Thr Ala Ala Thr
30 35 40

CCT GTC CAC CAC TGT CGC CCG CCC CAC AAT GCC TCC ACA GGG CCT TGG 252
Pro Val His His Cys Arg Pro Pro His Asn Ala Ser Thr Gly Pro Trp
45 50 55 60

GTG CTC CCC ATG GGC CCA AAT GGG AAG CCT GAG AGG TGC CTC CGT TTT 300
Val Leu Pro Met Gly Pro Asn Gly Lys Pro Glu Arg Cys Leu Arg Phe
65 70 75

GTA CAT CCG CCC AAT GCC AGC CTG CCC AAT GAC ACC CAG AGG GCC ATG 348
Val His Pro Pro Asn Ala Ser Leu Pro Asn Asp Thr Gln Arg Ala Met
80 85 90

GAG CCA TGC CTG GAT GGC TGG GTC TAC AAC AGC ACC AAG GAC TCC ATT 396
Glu Pro Cys Leu Asp Gly Trp Val Tyr Asn Ser Thr Lys Asp Ser Ile
95 100 105

GTG ACA GAG TGG GAC TTG GTG TGC AAC TCC AAC AAA CTG AAG GAG ATG 444
Val Thr Glu Trp Asp Leu Val Cys Asn Ser Asn Lys Leu Lys Glu Met
110 115 120

GCC CAG TCT ATC TTC ATG GCA GGT ATA CTG ATT GGA GGG CTC GTG CTT 492
Ala Gln Ser Ile Phe Met Ala Gly Ile Leu Ile Gly Gly Leu Val Leu
125 130 135 140

GGA GAC CTG TCT GAC AGG TTT GGC CGC AGG CCC ATC CTG ACC TGC AGC 540
Gly Asp Leu Ser Asp Arg Phe Gly Arg Arg Pro Ile Leu Thr Cys Ser
145 150 155

TAC CTG CTG CTG GCA GCC AGC GGC TCC GGT GCA GCC TTC AGC CCC ACC 588

Tyr Leu Leu Leu Ala Ala Ser Gly Ser Gly Ala Ala Phe Ser Pro Thr
160 165 170

TTC CCC ATC TAC ATG GTC TTC CGC TTC CTG TGT GGC TTT GGC ATC TCA 636
Phe Pro Ile Tyr Met Val Phe Arg Phe Leu Cys Gly Phe Gly Ile Ser
175 180 185

GGC ATT ACC CTG AGC ACC GTC ATC TTG AAT GTG GAA TGG GTG CCT ACC 684
Gly Ile Thr Leu Ser Thr Val Ile Leu Asn Val Glu Trp Val Pro Thr
190 195 200

CGG ATG CGG GCC ATC ATG TCG ACA GCA CTC GGG TAC TGC TAC ACC TTT 732
Arg Met Arg Ala Ile Met Ser Thr Ala Leu Gly Tyr Cys Tyr Thr Phe
205 210 215 220

GGC CAG TTC ATT CTG CCC GGC CTG GCC TAC GCC ATC CCC CAG TGG CGT 780
Gly Gln Phe Ile Leu Pro Gly Leu Ala Tyr Ala Ile Pro Gln Trp Arg
225 230 235

TGG CTG CAG TTA ACT GTG TCC ATT CCC TTC TTC GTC TTC TTC CTA TCA 828
Trp Leu Gln Leu Thr Val Ser Ile Pro Phe Phe Val Phe Phe Leu Ser
240 245 250

TCC TGG TGG ACA CCA GAG TCC ATA CGC TGG TTG GTC TTG TCT GGA AAG 876
Ser Trp Trp Thr Pro Glu Ser Ile Arg Trp Leu Val Leu Ser Gly Lys
255 260 265

TCC TCG AAG GCC CTG AAG ATA CTC CGG CGG GTT GGC TGT CTT CAA TGG 924
Ser Ser Lys Ala Leu Lys Ile Leu Arg Arg Val Gly Cys Leu Gln Trp
270 275 280

CAA GAA GGA AGA AGG AGA AAG CTC AGC TTG GAA GAG CTC AAA CTC AAC 972

Gln Glu Gly Arg Arg Arg Lys Leu Ser Leu Glu Glu Leu Lys Leu Asn
285 290 295 300

CTG CAG AAG GAG ATC TCC TTG GCC AAG GCC AAT TAC ACC GCA AGT GAC 1020
Leu Gln Lys Glu Ile Ser Leu Ala Lys Ala Asn Tyr Thr Ala Ser Asp
305 310 315

CTG TTC CGG ATA CCC ATG CTG CGC CGC ATG ACC TTC TGT CTT TCC CTG 1068
Leu Phe Arg Ile Pro Met Leu Arg Arg Met Thr Phe Cys Leu Ser Leu
320 325 330

GCC TGG TTT GCT ACC GGT TTT GCC TAC TAT AGT TTG GCT ATG GGT GTG 1116
Ala Trp Phe Ala Thr Gly Phe Ala Tyr Tyr Ser Leu Ala Met Gly Val
335 340 345

GAA GAA TTT GGA GTC AAC CTC TAC ATC CTC CAG ATC ATC TTT GGT GGG 1164
Glu Glu Phe Gly Val Asn Leu Tyr Ile Leu Gln Ile Ile Phe Gly Gly
350 355 360

GTC GAT GTC CCA GCC AAG TTC ATC ACC ATC CTC TCC TTA AGC TAC CTG 1212
Val Asp Val Pro Ala Lys Phe Ile Thr Ile Leu Ser Leu Ser Tyr Leu
365 370 375 380

GGC CGG CAT ACC ACT CAG GCC GCT GCC CTG CTC CTG GCA GGA GGG GCC 1260
Gly Arg His Thr Thr Gln Ala Ala Ala Leu Leu Leu Ala Gly Gly Ala
385 390 395

ATC TTG GCT CTC ACC TTT GTG CCC TTG GAC TTG CAG ACC GTG AGG ACA 1308
Ile Leu Ala Leu Thr Phe Val Pro Leu Asp Leu Gln Thr Val Arg Thr
400 405 410

GTA TTG GCT GTG TTT GGG AAG GGA TGC CTA TCC AGC TCC TTC AGC TGC 1356

Val Leu Ala Val Phe Gly Lys Gly Cys Leu Ser Ser Ser Phe Ser Cys
415 420 425

CTC TTC CTC TAC ACA AGT GAA TTA TAC CCC ACA GTC ATC AGG CAA ACA 1404
Leu Phe Leu Tyr Thr Ser Glu Leu Tyr Pro Thr Val Ile Arg Gln Thr
430 435 440

GGT ATG GGC GTA AGT AAC CTG TGG ACC CGC GTG GGA AGC ATG GTG TCC 1452
Gly Met Gly Val Ser Asn Leu Trp Thr Arg Val Gly Ser Met Val Ser
445 450 455 460

CCG CTG GTG AAA ATC ACG GGT GAG GTA CAG CCC TTC ATC CCC AAT ATC 1500
Pro Leu Val Lys Ile Thr Gly Glu Val Gln Pro Phe Ile Pro Asn Ile
465 470 475

ATC TAC GGG ATC ACC GCC CTC CTC GGG GGC AGT GCT GCC CTC TTC CTG 1548
Ile Tyr Gly Ile Thr Ala Leu Leu Gly Gly Ser Ala Ala Leu Phe Leu
480 485 490

CCT GAG ACC CTG AAT CAG CCC TTG CCA GAG ACT ATC GAA GAC CTG GAA 1596
Pro Glu Thr Leu Asn Gln Pro Leu Pro Glu Thr Ile Glu Asp Leu Glu
495 500 505

AAC TGG TCC CTG CGG GCA AAG AAG CCA AAG CAG GAG CCA GAG GTG GAA 1644
Asn Trp Ser Leu Arg Ala Lys Lys Pro Lys Gln Glu Pro Glu Val Glu
510 515 520

AAG GCC TCC CAG AGG ATC CTG TAC AGC CTC ACG GAC CAG GCC TGG GCT 1692
Lys Ala Ser Gln Arg Ile Leu Tyr Ser Leu Thr Asp Gln Ala Trp Ala
525 530 535 540

CCA GCT GAG GAC AAC GGA ACC CCC TTT CCC TGC CCT CCA GAG ACT GAT 1740

Pro Ala Glu Asp Asn Gly Thr Pro Phe Pro Cys Pro Pro Glu Thr Asp
545 550 555

CCT AGC CAG GCA CCT TAGGAGTATA GGGAGGCCCC ATATAGGTCC ATCCTCCTAG 1795
Pro Ser Gln Ala Pro
560

GATGAAGCCT TCTGAGAGCT TGGTGAAGGT GTCTCCATCA CCACCACCAG AGCCTCCTGC 1855

CCAGCCCTGG CCAGTTCAAA GGTTCAGCCA TCCCTGCCCT TGTTCTCCCT GCAACCCAGG 1915

CCCTGCCATT CTTCTGTCTA GCCCTTCCCC ACTGGCCACC TTCCCCCACT GTCCCGGTCC 1975

TCTTCCCCTG AGGTCCCCTG ATATCCCCTG GCTCAGTCCT AACAAGACTG AGTCTTAACA 2035

AGATGAGAAG TCCTCCCCTT CTTGCCTCCC ACACTTTTCT TTGATGGGAG GTTTCAATAA 2095

ACAGCGATAA GAACTCTAAA AAAAAAAAAA ACCGGAATTC 2135

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 561 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Thr Phe Ser Glu Ile Leu Asp Arg Val Gly Ser Met Gly His Phe

1 5 10 15

Gln Phe Leu His Val Ala Ile Leu Gly Leu Pro Ile Leu Asn Met Ala
20 25 30

Asn His Asn Leu Leu Gln Ile Phe Thr Ala Ala Thr Pro Val His His
35 40 45

Cys Arg Pro Pro His Asn Ala Ser Thr Gly Pro Trp Val Leu Pro Met
50 55 60

Gly Pro Asn Gly Lys Pro Glu Arg Cys Leu Arg Phe Val His Pro Pro
65 70 75 80

Asn Ala Ser Leu Pro Asn Asp Thr Gln Arg Ala Met Glu Pro Cys Leu
85 90 95

Asp Gly Trp Val Tyr Asn Ser Thr Lys Asp Ser Ile Val Thr Glu Trp
100 105 110

Asp Leu Val Cys Asn Ser Asn Lys Leu Lys Glu Met Ala Gln Ser Ile
115 120 125

Phe Met Ala Gly Ile Leu Ile Gly Gly Leu Val Leu Gly Asp Leu Ser
130 135 140

Asp Arg Phe Gly Arg Arg Pro Ile Leu Thr Cys Ser Tyr Leu Leu Leu
145 150 155 160

Ala Ala Ser Gly Ser Gly Ala Ala Phe Ser Pro Thr Phe Pro Ile Tyr
165 170 175

Met Val Phe Arg Phe Leu Cys Gly Phe Gly Ile Ser Gly Ile Thr Leu
180 185 190

Ser Thr Val Ile Leu Asn Val Glu Trp Val Pro Thr Arg Met Arg Ala
195 200 205

Ile Met Ser Thr Ala Leu Gly Tyr Cys Tyr Thr Phe Gly Gln Phe Ile
210 215 220

Leu Pro Gly Leu Ala Tyr Ala Ile Pro Gln Trp Arg Trp Leu Gln Leu
225 230 235 240

Thr Val Ser Ile Pro Phe Phe Val Phe Phe Leu Ser Ser Trp Trp Thr
245 250 255

Pro Glu Ser Ile Arg Trp Leu Val Leu Ser Gly Lys Ser Ser Lys Ala
260 265 270

Leu Lys Ile Leu Arg Arg Val Gly Cys Leu Gln Trp Gln Glu Gly Arg
275 280 285

Arg Arg Lys Leu Ser Leu Glu Glu Leu Lys Leu Asn Leu Gln Lys Glu
290 295 300

Ile Ser Leu Ala Lys Ala Asn Tyr Thr Ala Ser Asp Leu Phe Arg Ile
305 310 315 320

Pro Met Leu Arg Arg Met Thr Phe Cys Leu Ser Leu Ala Trp Phe Ala
325 330 335

Thr Gly Phe Ala Tyr Tyr Ser Leu Ala Met Gly Val Glu Glu Phe Gly
340 345 350

Val Asn Leu Tyr Ile Leu Gln Ile Ile Phe Gly Gly Val Asp Val Pro
355 360 365

Ala Lys Phe Ile Thr Ile Leu Ser Leu Ser Tyr Leu Gly Arg His Thr
370 375 380

Thr Gln Ala Ala Ala Leu Leu Leu Ala Gly Gly Ala Ile Leu Ala Leu
385 390 395 400

Thr Phe Val Pro Leu Asp Leu Gln Thr Val Arg Thr Val Leu Ala Val
405 410 415

Phe Gly Lys Gly Cys Leu Ser Ser Ser Phe Ser Cys Leu Phe Leu Tyr
420 425 430

Thr Ser Glu Leu Tyr Pro Thr Val Ile Arg Gln Thr Gly Met Gly Val
435 440 445

Ser Asn Leu Trp Thr Arg Val Gly Ser Met Val Ser Pro Leu Val Lys
450 455 460

Ile Thr Gly Glu Val Gln Pro Phe Ile Pro Asn Ile Ile Tyr Gly Ile
465 470 475 480

Thr Ala Leu Leu Gly Gly Ser Ala Ala Leu Phe Leu Pro Glu Thr Leu
485 490 495

Asn Gln Pro Leu Pro Glu Thr Ile Glu Asp Leu Glu Asn Trp Ser Leu
500 505 510

Arg Ala Lys Lys Pro Lys Gln Glu Pro Glu Val Glu Lys Ala Ser Gln
515 520 525

Arg Ile Leu Tyr Ser Leu Thr Asp Gln Ala Trp Ala Pro Ala Glu Asp
530 535 540

Asn Gly Thr Pro Phe Pro Cys Pro Pro Glu Thr Asp Pro Ser Gln Ala
545 550 555 560

Pro

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mus musculus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Asp Arg Phe Gly Arg Lys

1 5

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 2
(D) OTHER INFORMATION: /note= "Xaa at position 2 is Arg or Lys"

(ix) FEATURE:
(A) NAME/KEY: Region
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= "Xaa at position 5 is Arg or Lys"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Asp Xaa Xaa Arg Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 328 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAGTTCATCA CCATCCTCTC CTTAAGCTAC CTGGGCCGGC ATACCACTCA GGCCGCTGCC 60

TGCTCCTGGC AGGAGGGGCC ATCTTGGCTC TCACCTTTTG CCCTTGGACT TGCAGACCGT 120

GAGACAGTAT TGGCTGTGTT TGGGAAGGGA TGCCTATCCA GCTCCTTCAG CTGCCTCTTC 180

CTCTACACAA GTGAATTATA CCCCACAGTC ATCAGGCAAA CAGGTATGGG CGTAAGTAAC 240

CTGTGGACCC GCGTGGGAAG CATGGTGTCC CGCTGGTGAA AATCACGGGT GAGGTACAGC 300

CCTTCATCCC CAATATCATC TACGGGAT 328

CLAIMS

1. An isolated nucleic acid molecule

(a) which hybridizes under stringent conditions to a molecule consisting of the nucleic acid sequence of SEQ ID NO:1 or SEQ ID NO:7 and which codes for an osteoclast transporter molecule,

(b) nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code, and

(c) complements of (a) and (b).

2. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule hybridizes under stringent conditions to a molecule consisting of the nucleic acid sequence of SEQ ID NO:3.

3. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule consists essentially of SEQ ID NO:3.

4. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule consists essentially of SEQ ID NO:1.

5. An isolated nucleic acid molecule selected from the group consisting of (a) a unique fragment of SEQ ID NO:1 between 12 and 2101 nucleotides in length, (b) a unique fragment of nucleotides 1-2112 of SEQ ID NO:3 between 12 and 2111 nucleotides in length, (c) complements of "(a)" and (d) complements of "(b)".

6. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule is a unique fragment of SEQ ID NO:1 between 12 and 1974 nucleotides in length.

7. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule is a unique fragment of nucleotides 120-1733 of SEQ ID NO:1 between 12 and 1612 nucleotides in length.

8. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule is selected from the group consisting of a unique fragment of nucleotides 1-1179 of SEQ ID NO:3 between 12 and 1178 nucleotides in length and a unique fragment of nucleotides 1512-2112 of SEQ ID NO:3 between 12 and 599 nucleotides in length.

9. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule is a unique fragment of nucleotides 73-1758 of SEQ ID NO:3 between 12 and 1684 nucleotides in length.

10. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule is selected from the group consisting of a unique fragment of nucleotides 73-1179 of SEQ ID NO:3 between 12 and 1105 nucleotides in length and a unique fragment of nucleotides 1512-1758 of SEQ ID NO:3 between 12 and 245 nucleotides in length.

11. The isolated nucleic acid molecule of any of claims 5-10, wherein the isolated nucleic acid molecule is selected from the group consisting of a unique fragment of at least 14 contiguous nucleotides, a unique fragment of at least 15 contiguous nucleotides, a unique fragment of at least 16 contiguous nucleotides, a unique fragment of at least 17 contiguous nucleotides, a unique fragment of at least 18 contiguous nucleotides, a unique fragment of at least 20 contiguous nucleotides and a unique fragment of at least 22 contiguous nucleotides.

12. The isolated nucleic acid molecule of any of claims 5-10, wherein the isolated nucleic acid molecule consists of between 12 and 32 contiguous nucleotides.

13. A host cell transformed or transfected with an expression vector comprising the isolated nucleic acid molecule of any of claims 1, 2, 3 or 4 operably linked to a promoter.

14. An isolated polypeptide coded for by the isolated nucleic acid molecule of any of claims 1, 2, 3 or 4 .

15. An isolated polypeptide selectively binding a protein coded for by the isolated nucleic acid molecule of any of claims 1, 2, 3 or 4 .

16. The isolated polypeptide of claim 15, wherein the isolated polypeptide is an Fab or F(ab) fragment of an antibody.

17. The isolated polypeptide of claim 15 wherein the isolated polypeptide is a fragment of an antibody, the fragment including a CDR3 region selective for the protein.

18. The isolated polypeptide of claim 15, wherein the isolated polypeptide is a monoclonal antibody.

19. The isolated polypeptide of claim 15, 16 or 17, wherein the isolated polypeptide selectively binds an extracellular portion of the protein.

20. A method for decreasing osteoclast activity in a subject comprising administering to a subject in need of such treatment an agent that selectively binds to an isolated nucleic acid molecule of claim 1 or an expression product thereof, in an amount effective to decrease osteoclast activity in said subject.

21. A method as claimed in claim 20, wherein the agent is a modified nucleic acid.

22. A method as claimed in claim 20, wherein the agent is a polypeptide.

23. The use of an agent that selectively binds to an isolated nucleic acid molecule of claim 1, or an expression product thereof, in the preparation of a medicament.

24. The use as claimed in claim 23, wherein the agent is a modified nucleic acid.

25. The use as claimed in claim 23, wherein the agent is a polypeptide.

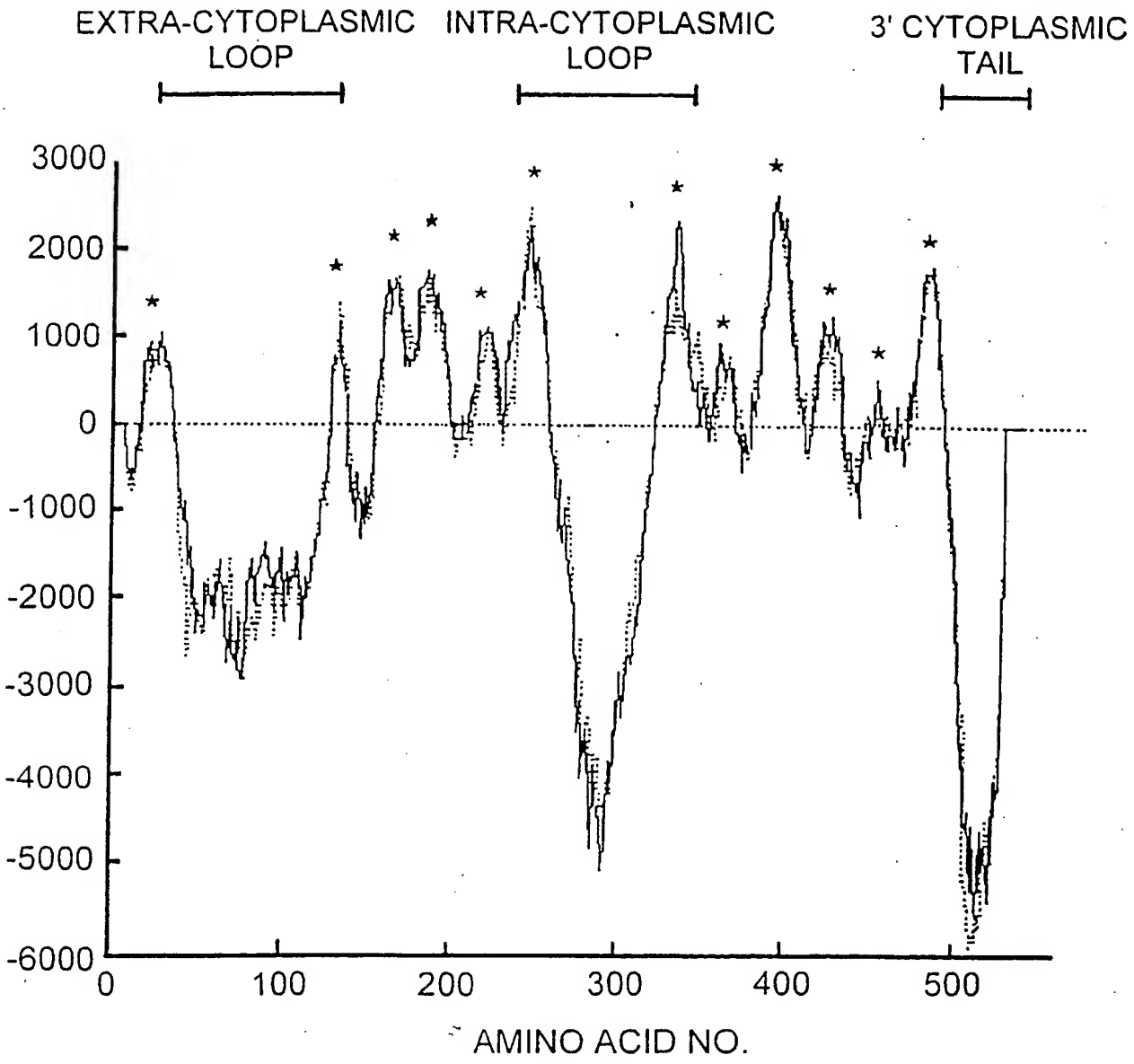

EXTRA-CYTOPLASMIC LOOP
INTRA-CYTOPLASMIC LOOP
3' CYTOPLASMIC TAIL
3000
2000
1000
0
-1000
-2000
-3000
-4000
-5000
-6000
0
100
200
300
400
500
AMINO ACID NO.

Fig. 1

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 97/07856

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/705 C07K16/28 C12N15/11 A61K39/00
A61K31/70 C12N1/21 C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EMBL Database entry Hs345165; Accession number H20345; 08.07.1995; Hillier L. et al.: 'The WashU-Merck EST project.' XP002038818 cited in the application see abstract ---	5,8-12
X	EMBL Database entry Hs581185; Accession number H41581; 05.08.1995; Hillier L. et al.: 'The WashU-Merck EST project.' XP002038819 see abstract ---	5,8,11, 12

-/--

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 August 1997	16.09.97

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Mandl, B

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EMBL Database entry Mm52842; Accession number U52842; 07.05.1996; Lopez-Nieto C.E.: 'Molecular cloning and characterization of a novel member of the organic cation transporter family.' XP002038820 see abstract	1-14
A	TIBS, vol. 18, 1993, pages 13-20, XP002038815 MARGER M.D. AND SAIER M.H.: "A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport." cited in the application see the whole document	1-14
A	NATURE, vol. 372, 8 December 1994, pages 549-552, XP002038816 GRÜNDEMANN D. ET AL.: "Drug excretion mediated by a new prototype of polyspecific transporter." cited in the application see the whole document	1-14
P,X	THE AMERICAN JOURNAL OF HUMAN GENETICS, vol. 59, no. 4, October 1996, page A280 XP002038817 BEIER D.R. ET AL.: "Absence of a novel putative transporter causes osteopetrosis in the osteosclerosis (oc) mouse." see abstract	1-25
P,X	EMBL Database entry MMAA8584; Accession number AA108584; 06.11.1996; Marra M. et al.: 'The WashU-HHMI mouse EST project.' XP002038821 see abstract	1-14

INTERNATIONAL SEARCH REPORT

PCT/US 97/07856

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Please see Further Information sheet enclosed.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/210

Although claims 20-22 are directed to a method of treatment of (diagnostic method practised on) the human/animal body the search has been carried out and based on the alleged effects of the compound/composition.